OPEN ACCESS

University of BRISTOL

Zhao, G., Bates, P. D., & Neal, J. (2020). The impact of dams on design floods in the Conterminous US. *Water Resources Research*, *56*(3), [e2019WR025380]. https://doi.org/10.1029/2019WR025380

Publisher's PDF, also known as Version of record

Link to published version (if available):
10.1029/2019WR025380

Link to publication record in Explore Bristol Research
PDF-document

This is the final published version of the article (version of record). It first appeared online via Wiley at https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR025380. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR025380

The Impact of Dams on Design Floods in the Conterminous US

Gang Zhao[1], Paul Bates[1,2], and Jeffrey Neal[1,2]

[1]School of Geographical Sciences, University of Bristol, Bristol, UK, [2]Fathom, Engine Shed, Station Approach, Bristol, UK

Key Points:
- A method is proposed to analyze the flood attenuation effect of dams in ungauged areas using publicly available data
- Based on nine factors available in standard dam databases the flood attenuation effect could be predicted with a Nash-Sutcliffe coefficient of 0.88
- Two case studies are used to test the reliability of the proposed method for simulating flooding downstream of dams

Supporting Information:
- Supporting Information S1
- Data Set S1
- Data Set S2

Correspondence to:
G. Zhao,
gang.zhao@bristol.ac.uk

Citation:
Zhao, G., Bates, P., & Neal, J. (2020). The Impact of Dams on Design Floods in the Conterminous US. *Water Resources Research*, *56*, e2019WR025380. https://doi.org/10.1029/2019WR025380

Received 16 APR 2019
Accepted 7 MAR 2020
Accepted article online 12 MAR 2020

Abstract We propose a method to describe the impact of dams on design floods for ungauged areas and validate the method over the conterminous US (CONUS). A Random Forest (RF) model was chosen to capture the relationship between the change in 100-year return period flow up- and downstream of different dams and dam parameters available in the Global Reservoir and Dam (GRanD) database. The results showed that: (1) the RF model showed a greater accuracy in terms of Nash Sutcliffe efficiency coefficient (0.92 in training and 0.88 in testing) than a benchmark Multiple Linear Regression model (0.68 in training and 0.61 in testing); (2) Dam inflow, upstream catchment area, and long-term average discharge at reservoir location were the three most important factors for dam outflow; (3) flood attenuation effect indices (FAI) for >1400 dams over the CONUS were derived with the proposed method. To further validate the accuracy of the FAI, a new module considering flood attenuation effects was developed for the LISFLOOD-FP hydrodynamic model and two dams in the CONUS were selected to compare simulated flooded area against Federal Emergency Management Agency flood hazard maps. The result showed that the overestimation in flood hazard maps caused by not taking dams into account can be significantly corrected using the FAI and the enhanced LISFLOOD-FP model. We conclude that the proposed methodology is a valid approach to describe the impact of dams on design floods, thereby improving the accuracy of flood hazard maps, especially in ungauged areas.

Plain Language Summary Dams have significantly changed hydrological processes, and this needs to be considered in flood hazard mapping. However, traditional flood hazard maps at national scales cannot accurately describe flooding downstream of dams because of the lack of detailed knowledge of reservoir operating rules. This research proposed a machine learning based approach to describe the impact of dams on downstream design floods using publicly available data. This approach has been successfully demonstrated over the continental US and could be easily extended globally. Furthermore, the proposed method was coupled with an enhanced hydraulic model and applied for flood hazard mapping in two case studies in the continental US. Compared with Federal Emergency Management Agency flood hazard maps, there was an obvious overestimation of flood extent downstream of each dam when reservoir operations are ignored. These overestimations can be significantly corrected using the proposed coupled model framework.

1. Introduction

Flooding is regarded as one of the most destructive natural disasters and causes serious loss of life and economic damage worldwide (Jha et al., 2012). From 1972 to 2006, there were 531 US floods designated as catastrophes where losses exceeded $176 million per event (Changnon, 2008) and flood risk is predicted to increase given climate and socio-economic change (Ashley et al., 2005; Ntelekos et al., 2010; Wing et al., 2018). To mitigate flood hazard, various types of structural measures, such as dams, have been built along rivers. According to the World Register of Dams (ICOLD, 2018), more than 58,000 large dams (dam height > 15 m) have been constructed globally. The number of dams is still increasing each year, especially in developing countries (Gleick, 2012; Linnerooth-Bayer & Mechler, 2015). The basic functions of dams are to suppress downstream flooding and to store water for irrigation, human consumption or electricity generation. Dams have significantly changed hydrological processes, and this needs to be considered in flood hazard mapping. The purpose of this research was therefore to develop a globally applicable method to infer the impact of dams on downstream flood hazard.

With recent increases in computing power and the development of remote sensing technology, large-scale (regional to global) flood inundation modeling has become possible over the past decade (Hirabayashi

©2020. American Geophysical Union. All Rights Reserved.

et al., 2013; Pappenberger et al., 2012; Sampson et al., 2015; Ward et al., 2013). The flood hazard maps produced by these models fill significant gaps in ungauged areas and can provide valuable information for flood risk managers. Large-scale flood inundation models split the catchment up into several reaches and the flood hazard of each reach is simulated by computing the inundation that results from particular design floods (Sampson et al., 2015). Thus, the impact of dams on flood hazard can be reflected by the change in the magnitude of design floods downstream of reservoirs. However, dam characteristics are complex, and the operation process of each dam tries to balance multiple objectives such as flood control, water supply, hydroelectric power generation or other environmental objectives (Graf, 1999). Therefore, the impact of dams on flood hazard will alter depending on different reservoir operating purposes. However, to ensure the safety of the dam structure and downstream areas, the water level of dams is usually not allowed to exceed the flood control pool level before flood events (Valdes & Marco, 1995). It is reasonable to assume that dam operating rules simplify considerably during extreme floods (return periods of 50 years or greater). In such situations, dams are very likely to be operated to maximize flood control. As a result, we select a low frequency design flood (100-year return period) to analyze in this study.

The impact of individual or multiple dams on flow characteristics has been well analyzed in case studies worldwide (eg. Assani et al., 2006; Erskine et al., 1999; Fortier et al., 2011; Mei et al., 2017; Richter et al., 1998). Studies have adopted the ratio of storage capacity to mean annual inflow as an index to describe the impact of reservoir operations on annual streamflow (Friesen et al., 2005; McMahon et al., 2007). For flood flow characteristics, Assani et al. (2006) compared the impacts of dams on annual maximum flow characteristics between natural and regulated rivers in Quebec, Canada, and revealed that the type of flow regime and watershed size were the main factors that affected the amplitude of flow changes. However, all the above analyses have been based on observed discharge at regional scales and cannot easily be extended to large scale studies where many catchments will be ungauged. To address this issue, FitzHugh and Vogel (2011) assessed the impact of dams on flood flows in the conterminous United States (CONUS) by developing regression models between median annual 1-day maximum flow and several watershed characteristics in different hydrologic units. FitzHugh and Vogel's research covered 78% of the CONUS and 81% of the developed regressions were considered to be representative enough for analysis. One problem identified by FitzHugh and Vogel (2011) was that it would be dangerous to extrapolate the results developed for a particular hydrologic unit to other ungauged areas. Also, flood hazard mapping is usually analyzed based on return period floods which were not discussed in FitzHugh and Vogel's research.

The preferred method to calculate dam regulated flow under different return periods is to represent local reservoir operation rules in hydrological models (eg. Ahmad et al., 2014; Guo et al., 2011; Pan et al., 2015; Yang et al., 2017). These operation rules are usually commercially sensitive and need detailed dam and reservoir parameters that are difficult to obtain globally. Although several studies have incorporated simple reservoir operation rules into large-scale hydrological models to simulate the impact of dams on long-term (monthly or seasonal) streamflow (Biemans et al., 2011; Döll et al., 2009; Hanasaki et al., 2006; Zhou et al., 2016), simulation of design floods has not been addressed by any of these studies. To date, one of the few studies to have looked at this is that of Zajac et al. (2017) who analyzed the impact of 667 large dams on design floods using a global application of the LISFLOOD hydrological model (note this scheme is different from the LISFLOOD-FP model mentioned earlier which is a two-dimensional hydrodynamic code) and a simple dam module. In Zajac et al.'s research, the dam outflow is calculated according to four sets of rules in the dam module of the LISFLOOD model. Because of the shortage of consistent, global operational records for dams, the reservoir parameters in the dam module were set by fraction of total dam capacity and percentiles of naturalized daily streamflow in Zajac et al.'s study. By adopting global sensitivity and uncertainty analysis methods, Zajac et al. (2017) showed considerable uncertainty in the downstream discharge of dams as a result of variations in these parameters. Overall, research concerning the impact of dams on design floods is therefore still inadequate.

Typically, the impact of dams on the magnitude of design floods is analyzed by looking for changes in the flood frequency curve before and after reservoir operations. One difficult task during analysis is to estimate the total dam inflow under different return periods because observed streamflow on all reservoir upstream tributaries is frequently lacking. In this study, the index-flood method proposed by Smith et al. (2015) was adopted to derive design flood values for the upstream tributaries of each dam. Index-flood methods were first described by Dalrymple (1960) and have long been regarded as a mature regional flood frequency

analysis (RFFA) approach for ungauged areas (Dalrymple, 1960). The basic assumptions of the index-flood method are that the flood frequency curve at different sites in one region is the same except for one scale parameter (known as the index-flood) (Bocchiola et al., 2003; Dalrymple, 1960); and this relationship from data-rich regions can be transferred to data poor ones with similar properties. A rich literature now exists on the estimation of design floods via index-flood methods at regional scales (Bocchiola et al., 2003; Kjeldsen et al., 2002; Kjeldsen & Jones, 2006) and Smith et al. (2015) for the first time extended the method to global coverage based on Global Runoff Data Centre (GRDC) database. The Smith et al.'s (2015) research adopted precipitation, slope and catchment area to estimate return period flows of reaches and has shown skill in characterizing natural flood behaviors at the global scale. This method has been successfully used to generate design floods for global flood hazard modeling for the United Kingdom, Canada, and CONUS (Sampson et al., 2015; Wing et al., 2017). However, the impact of dams on design floods was not considered in these researches and this will cause overestimation of flood hazard downstream of dams. To address this problem, we here attempt to develop a quantitative method to evaluate dam effects on design floods for ungauged areas based on publicly available datasets. This method can also provide boundary conditions for hydraulic models and can give a more comprehensive analysis of the impact of dams on flood hazard.

The method described herein quantifies the impact of dams on the design flood (100-year return period) for ungauged areas and validates the method over the CONUS. 360 dams with downstream observed stream gages in the CONUS were selected for model development. The inflow and outflow of these dams were estimated by index-flood methods and flood frequency analysis of stream gage data respectively for the 100-year return period flow. A Random Forest (RF) model was chosen to capture the relationship between the change in 100-year return period flow up- and downstream of different dams and dam parameters available in the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011), and the RF performance was compared with a benchmark Multiple Linear Regression (MLR) model. The proposed method allows us to quantitatively evaluate the change in design flood with and without the effects of reservoir operation for ungauged areas. The flood hazard considering dam effects was mapped in two case studies by coupling these estimates with an enhanced LISFLOOD-FP hydrodynamic model at 30-meter spatial resolution. The simulated flood extent with and without taking the effect of reservoir operations into account was validated against FEMA maps of the corresponding flood hazard.

2. Dams and Data Description

In this study, 1403 dams in the GRanD database located on the river network of the CONUS were selected for analysis. The river network is extracted from a 1 km resolution HydroSheds flow accumulation map (Lehner et al., 2008) for catchments exceeding a threshold area of 50 km^2. Among all dams, 360 representative dams were selected for model implementation based on three criteria. Firstly, the dam should be included in the GRanD. GRanD is a public dam and reservoir data set which provides multiple attributes of dams and reservoirs worldwide. Secondly, there should be one streamflow gage in the immediate downstream vicinity of the dam to represent the reservoir outflow. Thirdly, this downstream gage should have long time series of observed daily streamflow (more than 20 years) after dam construction. The location and time series of observed daily streamflow of these stream gages were downloaded from the U.S. Geological Survey website (waterdata.usgs.gov). The location of the selected dams is shown in Figure 1 of which 70% and 30% were randomly selected for regression model training and testing, respectively.

3. Methods Description

3.1. Framework

The framework of this study is shown in Figure 2 and can be divided into four parts. Part (1) describes the dam inflow and outflow estimation based on gauge data. Part (2) describes the processes of statistical model development and comparison. In this part, relationships between dam inflow, outflow, and dam parameters are built based on RF and MLR models respectively. Part (3) describes the dam in/outflow estimation in ungauged areas. The dam inflow was estimated by the calibrated index-flood method in Part (1) and the dam outflow is derived from the validated RF model in Part (2). Part (4) explains how the results are applied to flood hazard mapping in two case studies and the simulated flood hazard map is compared with a public

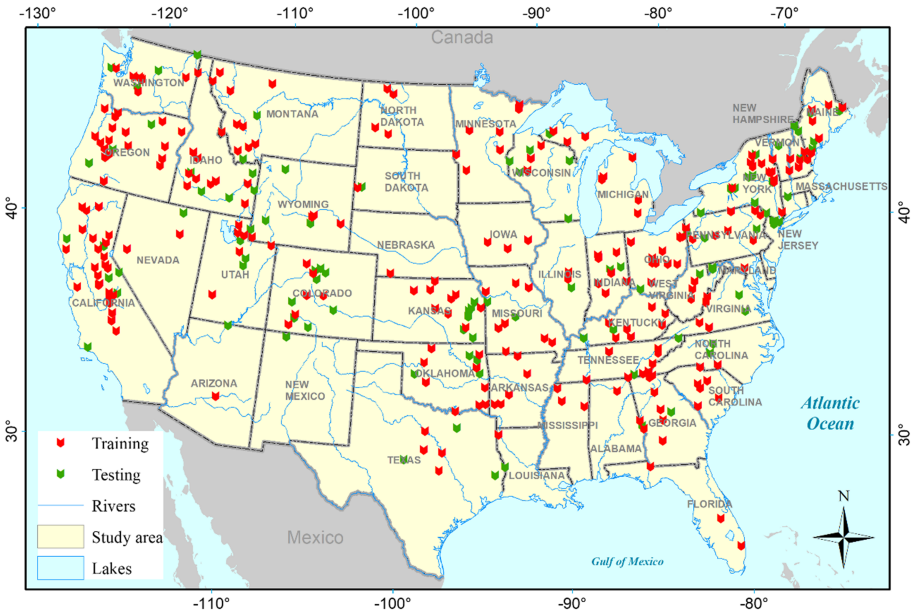

Figure 1. Study areas and location of dams in training and testing periods.

flood hazard map in the CONUS. The methods used in parts (1–3) are described in this section and the details of case studies and the enhanced hydraulic model are described in Section 6 "Case studies".

3.2. Design Flood Estimation

When experiencing extreme floods, we assume that the main purpose of dams is flood mitigation. The attenuation effect on extreme floods is mainly controlled by reservoir inflows and the dam parameters. The relationship between the dam inflow and outflow under 100-year return period conditions can be formulated as equation 1.

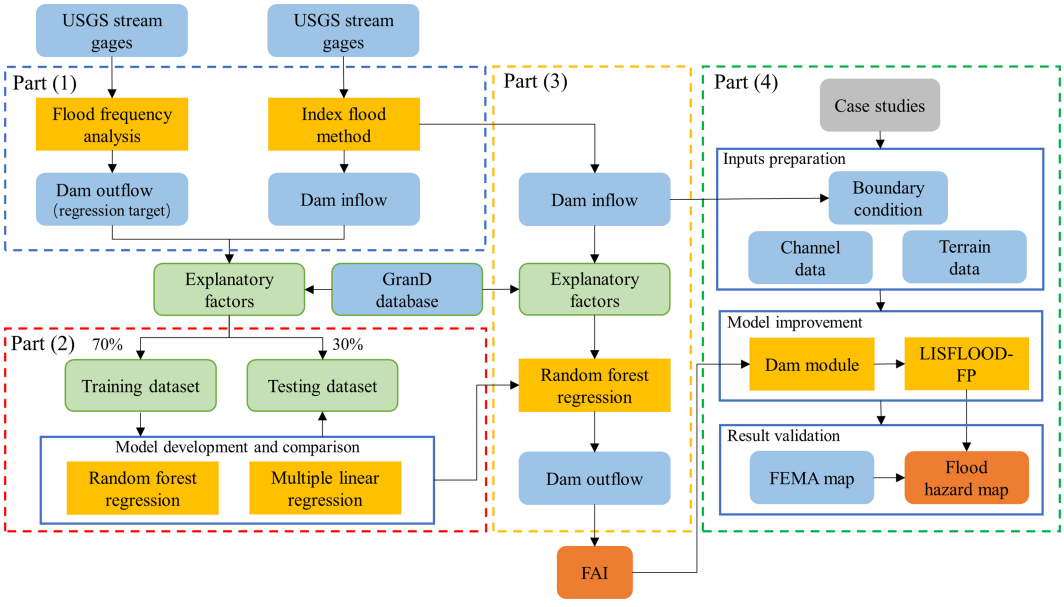

Figure 2. Framework of this study: (1) dam in/outflow estimation in gauged areas; (2) regression model development and comparison; (3) dam in/outflow estimation in ungauged areas (4) application of FAI in case studies.

$$QO = f(QI, p_n) \tag{1}$$

Where: QO is the dam outflow for a 100-year return period (1% annual probability); QI is the dam inflow for a 100-year return period; $f(-)$ is an unknown attenuation function; p_n (n = 1,2,3, ...,N) are dam or reservoir parameters and N is the number of parameters.

To better couple with the LISFLOOD-FP hydrodynamic model, a flood attenuation index (FAI) was introduced to describe the design flood change after reservoir operations. The definition of FAI is formulated as in equation 2 and it reflects the ratio of dam outflow to inflow at the flood peak.

$$FAI = \frac{QO}{QI} \tag{2}$$

Where: FAI is flood attenuation effect index for the 100-year return period flood.

The flood attenuation effect increases with the decline of FAI and the FAI is categorized into five classes (lowest: >0.80, low: 0.61–0.80, medium: 0.41–0.60, high: 0.20–0.60 and highest: <0.20). In this research, the CONUS is delineated into 199 subregions according to the USGS Hydrologic Unit Maps (Seaber et al., 1987). The mean value of FAI in the subregion was applied to describe the flood attenuation effect for each subregion.

3.2.1. Flood Frequency Analysis

The QO of each dam was calculated by flood frequency analysis based on stream gages data. Time series of annual maximum peak flows downstream of the dam after dam construction were extracted and the Generalized Extreme Value (GEV) distribution was selected as the flood frequency curve to fit the distribution of annual maximum peak flows. The cumulative distribution function of GEV can be described as in equation 3. The QO of each dam was interpolated from the flood frequency curve at 0.01 frequency.

$$\mathrm{GEV}(x|p) = \exp\left[-\left(1-\frac{b}{a}(x-c)\right)^{\frac{1}{b}}\right] \tag{3}$$

Where: x is time series of annual maximum peak flow; a is the scale parameter; b is the shape parameter, and c is the location parameter of the GEV distribution.

3.2.2. Index Flood Method

It is difficult to calculate the dam inflow directly based on observed streamflow data because for many dams not all upstream tributaries have streamflow gages. In this study, the design flood of upstream tributaries was calculated by the index-flood method proposed by Smith et al. (2015). This method can estimate the natural design flood in ungauged areas at global scales based on the assumption that there is a certain correlation between the index flood and climatic and physiographic factors of basins. Specifically, in this method the terrestrial land surface was first divided into several homogeneous subareas by a two-stage clustering model. In the clustering model, the subareas were preliminarily defined by Ward's algorithm (Ward, 1963) and then further refined using a K-means model (Lloyd, 1982). Secondly, three factors from publicly available datasets (annual precipitation from Hijmans et al., 2005, slope and catchment areas from Lehner et al., 2008), were selected as catchment descriptors. The index-flood (mean of annual peak flow) in each subarea was derived by fitting a power-form function as in equation 4 with the observed discharge data taken from the GRDC database. The calibrated equation 4 in each subarea can be used to estimate ungauged MAF for that subarea with the same catchment descriptors. Thirdly, the average ratio of observed design flood to the observed index-flood (r) was calculated in each subarea. Finally, the ungauged design flood can be estimated by the ratio (r) in the subarea and the estimated index-flood as in equation 5. The detailed procedure and the database for this method is described in detail in Smith et al. (2015). This method has been successfully used for flood inundation mapping in CONUS with a hit rate (the ratio of correct estimated area to flooded area of FEMA maps) of 86% (Wing et al., 2017).

$$MAF = K_1 \times A^{K2} \times P^{K3} \times S^{K4} \tag{4}$$

$$QI = \mathrm{MAF} \times r \tag{5}$$

Where: MAF is the mean of annual peak flows (known as the index-flood), A is catchment area, P is the annual precipitation, S is the average slope of the upstream catchment and K_n (n = 1,2,3 and 4) are parameters derived by Smith et al. (2015) which vary in each subarea; QI is the design flood in ungauged areas; r is the average ratio of observed design flood to the observed index-flood in each subarea.

3.3. RF Model Description

It may be difficult to describe the attenuation function in equation 1 with a simple linear relationship because of the complexity of hydrological process in a basin. In these circumstances, it is important to find a robust regression model that can capture non-linear relationships. Random Forest regression has gained attention recently for such a task because of its good performance in terms of noise, outliers and overfitting (Breiman (2002); Pang et al., 2017; Zhao et al. (2018); Woznicki et al., 2019) and was therefore used in this study. A Multiple Linear Regression (MLR) model was selected a benchmark statistical approach and is described in Supplementary Information 1. The model structure and parameter selection process of the RF model are described in Supplementary Information 2.

One advantage of RF is that it has an inbuilt cross-validation procedure during the training process that is achieved by analyzing out-of-bag (OOB) samples. Once a regression tree model is developed by bootstrap samples, the OOB samples (samples not in the bootstrap samples) can be used to test its accuracy. The whole learning error of RF is calculated by averaging the prediction error (mean square error) of each individual tree with corresponding OOB samples. Another advantage of the RF approach is that the importance of each inputted factor can be evaluated with the OOB error. An RF model estimates the importance of each factor by comparing how much prediction error increases when OOB samples for that factor are permuted while all others are unchanged. RF has been successfully applied in flood hazard assessment and has outperformed other machine intelligence methods such as Artificial Neural Networks and Support Vector Machines either at regional or national scale (Wang et al., 2015; Zhao et al., 2018).

3.4. Factor Selection

There are many dam and reservoir parameters contained in the GRanD database. Thus, a factor selection method is needed before model development to avoid the curse of dimensionality and reduce overfitting. The correlation-based feature selection (CFS) method (Hall, 1999) was adopted to select explanatory factors from all dam and reservoir parameters in the GranD database. This method assumes that irrelevant factors show a low correlation with the learning target and therefore can be ignored by the algorithm. The Pearson correlation between each factor and the target (*CFT*) in this research was evaluated as in equation 6. In this research, the factors in GRanD database are ignored when the *CFT* value was less than 0.05 leading to the selection of 9 factors for model development.

$$CFT = \frac{\sum_{n=1}^{N}\left(F^n-\overline{F}\right)\left(T^n-\overline{T}\right)}{\sqrt{\sum_{n=1}^{N}\left(F^n-\overline{F}\right)^2}\sqrt{\sum_{n=1}^{N}\left(T^n-\overline{T}\right)^2}} \tag{6}$$

Where F is the factor in the GRanD database; T is the model target (Dam outflow); N is the number of selected dams; *CFT* is the Pearson correlation coefficient between F and T. $\overline{F}$ and $\overline{T}$ are the average values of F and T respectively.

The factor importance can be evaluated by the *CFT* value in the CFS method according to equation 7:

$$FI_i = \frac{|CFT_i|}{\sum_{i=1}^{K}|CFT_i|} \tag{7}$$

Where: FI_i is the factor importance of F_i, $i = 1, 2, \ldots, K$; K is the total number of factors. $|CFT_i|$ is the absolute CFT_i value of F_i in the CFS method.

3.5. Evaluation Indices

The Nash Sutcliffe efficiency coefficient (*NSE*) and Pearson correlation coefficient (PCC) were used to evaluate the regression model performance for the training and testing datasets. These evaluation measures are described in equations 8 and 9.

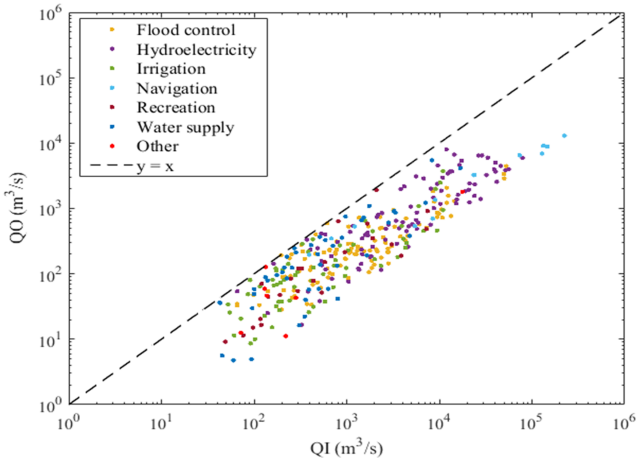

Figure 3. Dam inflow (QI) discharge estimated by the index-flood method versus dam outflow (QO) estimated with a frequency analysis method.

$$NSE = 1 - \frac{\sum_{n=1}^{N}\left(Q_s^n - Q_o^n\right)^2}{\sum_{n=1}^{N}\left(Q_o^n - \bar{Q}_o\right)^2} \tag{8}$$

$$PCC = \frac{\sum_{n=1}^{N}\left(Q_s^n - \overline{Q}_s\right)\left(Q_o^n - \overline{Q}_o\right)}{\sqrt{\sum_{n=1}^{N}\left(Q_s^n - \overline{Q}_s\right)^2}\sqrt{\sum_{n=1}^{N}\left(Q_o^n - \overline{Q}_o\right)^2}} \tag{9}$$

Where Q_s^n is the simulated peak flow of dam n (m^3/s); Q_o^n is the observed peak flow of dam n (m^3/s); $\bar{Q}_s$ is the mean value of simulated peak flow of all dams (m^3/s); $\bar{Q}_o$ is the mean value of observed peak flow of all dams (m^3/s) and N is the total number of dams.

Apart from *NSE* and PCC, two case studies were adopted to further test the effect of the attenuation index in flood hazard mapping by comparing the simulated flood extent using the new design discharge with public flood hazard maps developed by the Federal Emergency Management Agency (FEMA). To compare with the benchmark map, two widely used indices, the Critical Success Index (C) and Error Bias (E), were adopted in this study. These indices are described in equations 10 and 11.

$$C = \frac{M_1B_1}{M_1B_1 + M_0B_1 + M_1B_0} \tag{10}$$

$$E = \frac{M_1B_0}{M_0B_1} \tag{11}$$

Where M_1B_1 is the area that is inundated both in the modeled and benchmark map (m^2); M_1B_0 is the area that is inundated in modeled map but non-inundated in the benchmark (m^2); M_0B_1 is the area that is non-inundated in modeled map but inundated in benchmark (m^2); M_0B_0 is the area that is non-inundated both in the modeled and benchmark map (m^2);

4. Results

4.1. Explanatory Factors

Figure 3 shows the attenuation effect for dams with different purposes by plotting QI versus QO. A wide variety of attenuation effects was found. We found that QO and QI showed an obvious nonlinear relationship, however, no strong relationship was found between attenuation effects and dam purpose. This means that QO cannot be accurately described simply with QI and a knowledge of the use to which the dam is put. Instead, more explanatory factors need to be considered in model development and these were taken from the GRanD database.

After factor selection, a total number of nine explanatory factors were selected for RF model development. QO was calculated using flood frequency analysis of the stream gages (Section 3.2.1) and was regarded as the training target. QI was derived by index-flood method (Section 3.2.2) and the other eight explanatory factors were selected from the GranD database. The statistics of the explanatory factors and the target used are shown in Table 1.

4.2. Model Evaluation

Results from the RF model and the benchmarking MLR model are shown in Figure 4. We found that the RF method outperformed MLR both in the training and testing process. Specifically, RF had a better performance with NSE and PCC values of 0.92 and 0.97 during training and 0.88 and 0.92 during testing. MLR achieved NSE and PCC values of 0.68 and 0.87 in the training process and 0.61 and 0.78 in the testing process.

4.3. Factor Importance

Figure 5 shows the contribution of each factor to QO evaluated by the RF model and the CFS method respectively. As shown in Figure 5, dam inflow (QI), long-term average discharge (LD) and area of the upstream

Table 1
Statistics of Explanatory Factors and Target Data used in this Study

Factors/target	Name	Min	Max	Data source
F_1 (YC)	Year of construction	1835	1989	GranD database
F_2 (LG)	Length of dam (m)	5	11582	
F_3(AP)	Surface area of associated reservoir polygon (km^2)	0.3	599.6	
F_4 (AR)	Most reliable reported surface area of reservoir (km^2)	0.3	599.6	
F_5(AD)	Average depth of reservoir (m)	1.4	207.7	
F_6(LD)	Long-term (1961–90) average discharge at reservoir location (l/s)	1	1452733	
F_7(ES)	Elevation of reservoir surface in meters above sea level (m)	12	3135	
F_8(AC)	Area of upstream catchment draining into the reservoir (km^2)	90	278955	
F_9(QI)	Dam inflow under 100-year return period (m^3/s)	42	225190	Smith et al. (2015)
T (QO)	Dam outflow under 100-year return period (m^3/s)	4.7	12954	USGS stream gages

catchment (AC) are the three most important factors in both the RF model and CFS method. These factors mainly reflect the streamflow characteristics of the basin. The secondary set of important factors are the descriptors of the reservoir characteristics, including elevation (ES) and area of the reservoir surface (AP and AR). Average depth of reservoir (AD), Length of dam (LG) and year of dam construction (YC) are shown to be less important in both the RF and CFS methods.

The RF model performance was further tested by using different sets of model inputs and the results are shown in Table 2. We found that RF model performance improved with an increasing number of explanatory factors. The best model performances were found when adopting all explanatory factors for RF model

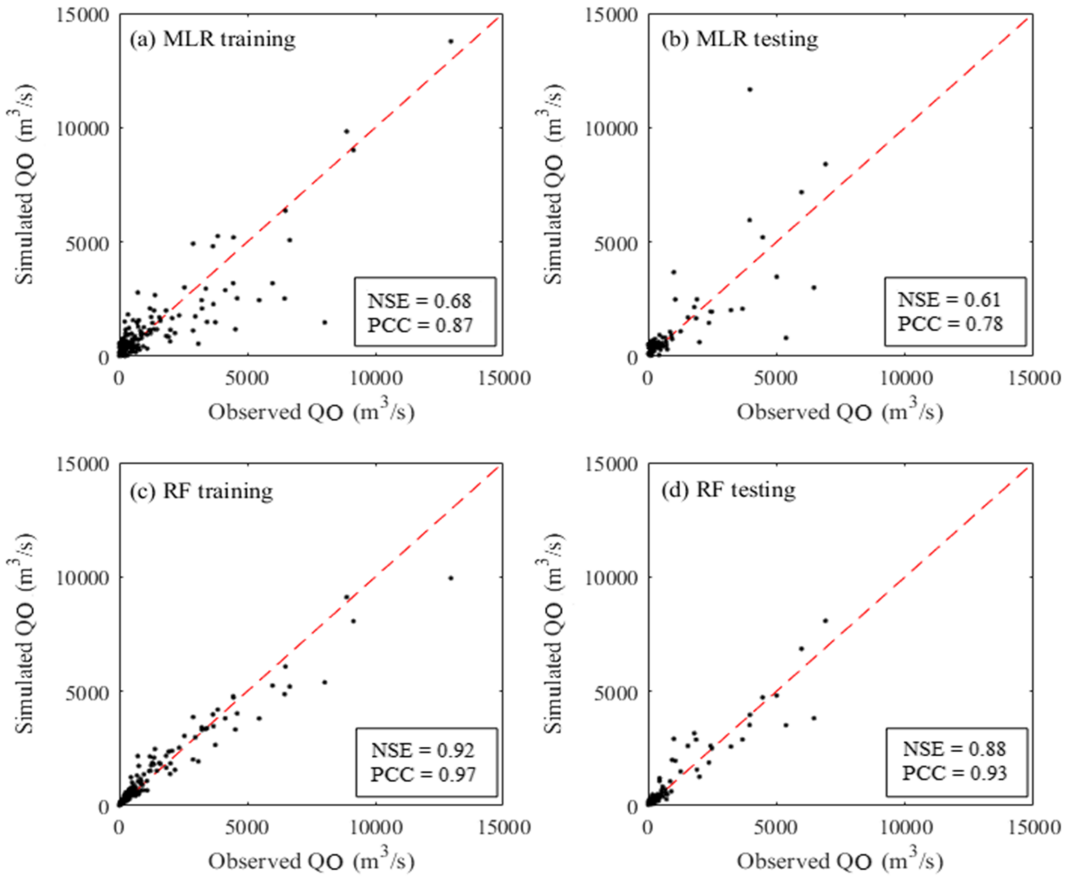

Figure 4. Performances of the RF method and benchmark MLR models in training and testing periods; (a and b, performances of MLR in training and testing; c and d, performances of RF in training and testing;)

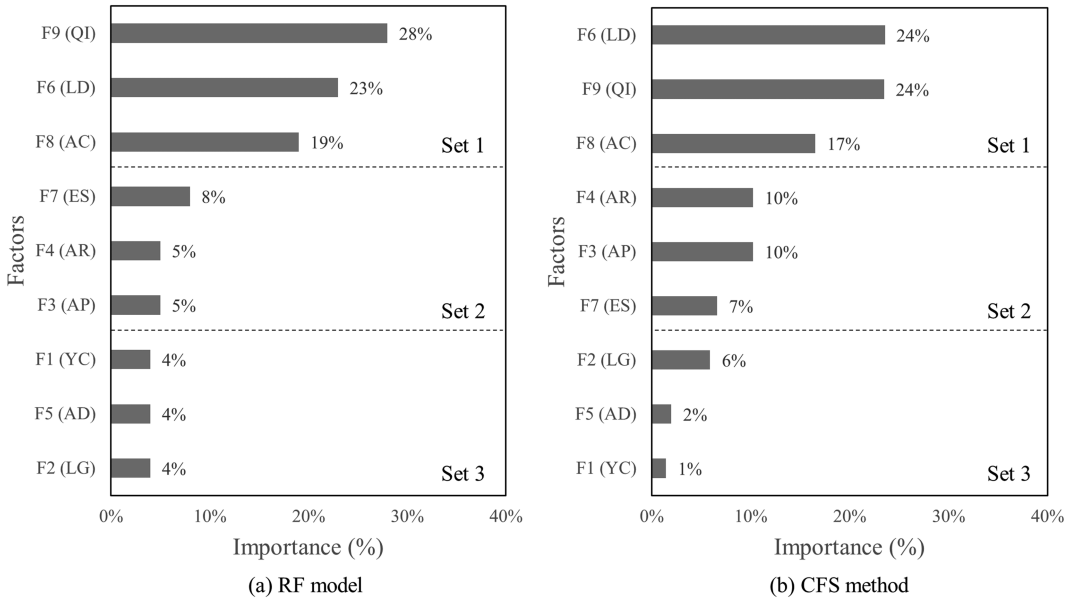

Figure 5. Factor importance evaluated by RF model a and CFS method b.

development, giving NSE values of 0.94 and 0.88 in the training and testing periods respectively. QO cannot be accurately predicted when only using QI as the model input (NSE < 0.6 both in the training and testing periods). The RF model still showed a satisfactory accuracy when using the parameters in the GRanD database for prediction. This revealed that this approach can be used in ungauged areas where the value of QI is not reliable.

4.4. FAI Results

The flood attenuation indices of 1403 dams in the CONUS were derived with the RF model and the distribution of FAI is shown in Figure 6 (a). As shown in Figure 6 (a), there is no obvious pattern or trend in FAI at the spatial scale of the CONUS mainly because the heterogeneity of the dam characteristics. About half of dams (47%) have high flood attenuation effects (FAI < 0.4) and 36% dams have a low flood attenuation effect (FAI > 0.6). The mean FAI value of 199 hydrologic units (HUs) in the CONUS is shown in Figure 6 (b). We found that 44 HUs experiences high attenuation effects and these are mainly located in the Pacific Northwest, the Lower Colorado basin, the lower Mississippi, Tennessee, the Great Lakes region and the middle reach of the Missouri. About 62% (124 of 199) of HUs in the CONUS have medium dam attenuation effects or larger. This analysis demonstrates how dams have significantly changed streamflow characteristics across the CONUS, and this should be considered in flood hazard modeling. 19 HUs have no data and this can be improved by collecting more dam records for future studies.

Table 2
The Impact of Factor Numbers on RF Model Performance

		Training		Testing	
No.	Factors	NSE	PCC	NSE	PCC
1	Set 1	0.92	0.96	0.83	0.91
2	Set 1 and 2	0.93	0.97	0.88	0.94
3	Set 1, 2, and 3	0.94	0.98	0.88	0.94
4	Only QI	0.51	0.87	0.57	0.90
5	All Sets without QI	0.92	0.97	0.86	0.93

5. Discussion on Discharge Uncertainty

The high fitting ability and prediction accuracy of RF revealed that the proposed approach can successfully be used in dam attenuation effect simulation with the available explanatory factors in the CONUS. Compared with the explanatory factors collected from the GRanD database directly, QI and QO were estimated by the index-flood method and a frequency analysis method respectively. The uncertainties during this estimation affect the accuracy of the RF model and derived FAI values. The impacts of QI and QO uncertainties on the results were quantitatively evaluated by randomly introducing different levels of bias as described in the Supplementary Information 3. The impacts of discharge uncertainty on RF model accuracy are presented in Table 3.

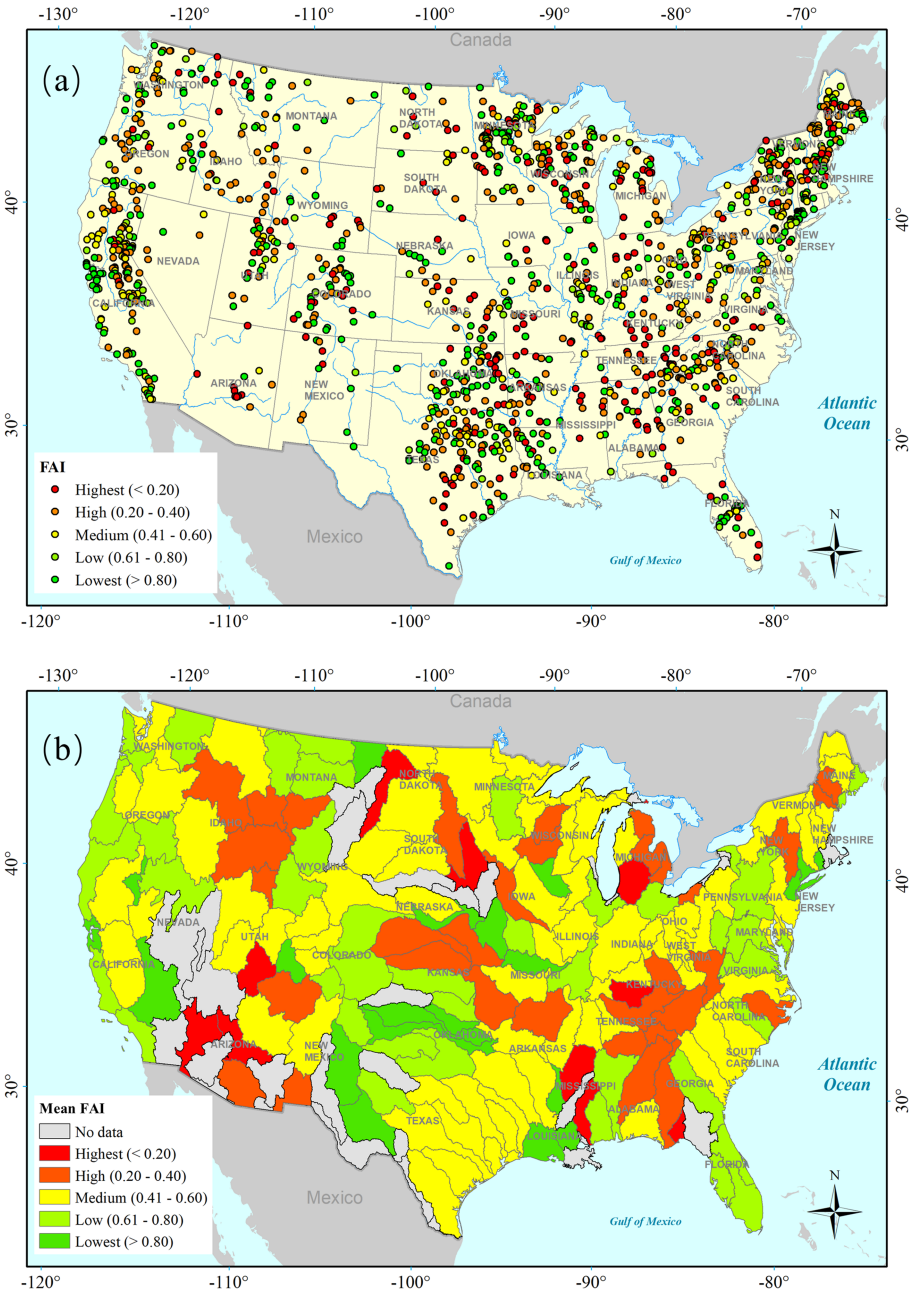

Figure 6. FAI distribution a and mean value of FAI in each hydrologic unit b based on RF model in the CONUS.

As shown in Table 3, the range of NSE values reflects the impact of dam in/outflow uncertainty on RF model accuracy. We found that model uncertainty increased with a larger bias adding to both QI and QO. Compared with QI, RF model accuracy is more sensitive to QO. The RF model can provide a reliable prediction (NSE > 0.85 in the testing period) when adding ±20% bias to QI. However, the NSE value of the RF model during testing could reduce to 0.713 when adding ±20% bias to QO.

Figure 7 displays the maximum and minimum value of FAI from 1000 RF models when adding±20% bias to QI and QO respectively. Most FAI values can be estimated within ±20% bias which revealed that the RF model will not exacerbate the effect of discharge uncertainty on the FAI value. This is mainly because the RF model is an ensemble result from a large number of tree models and each tree model is developed using a selected subset of the training dataset. Thus, large error inputs will not significantly change the RF model

Table 3
The Impact of Dam in/Outflow Uncertainty on RF Model Accuracy

Bias levels	QI						QO					
	NSE in training			NSE in testing			NSE in training			NSE in testing		
	Max	Min	Range	Max	Min	Range	Max	Min	Range	Max	Min	Range
±10%	0.944	0.933	0.011	0.899	0.871	0.029	0.949	0.927	0.021	0.915	0.834	0.081
±15%	0.946	0.933	0.013	0.901	0.866	0.036	0.949	0.923	0.026	0.925	0.817	0.107
±20%	0.946	0.933	0.013	0.903	0.853	0.049	0.952	0.917	0.035	0.937	0.713	0.223

structure. As shown in Figure 7, the range of FAI for QI is wider than for QO which reveals that QI uncertainty has a lager impact on FAI results. Although some individual large error points were found both under QI and QO uncertainties, the proposed method could nevertheless provide valuable information about flood attention effects for decision makers at the national scale.

6. Case Studies

In this section, the proposed approach for determining FAI was coupled with an enhanced hydraulic model and applied for flood hazard mapping in two case studies in the CONUS. For validation, the model outputs are compared with Federal Emergency Management Agency (FEMA) 100-year return period flood hazard maps. FEMA maps are the official source of flood hazard data in the US and are produced to nationally agreed quality standards. The FEMA map is produced from individual reach-scale flood hazard studies undertaken with different hydrological datasets and hydraulic models. The FEMA map describes the flood hazard in main river reaches only and covers about 66% of the total area of CONUS (Wing et al., 2017). The flood hazard for the two case studies is simulated by a widely used high-resolution flood hazard model framework (LISFLOOD-FP: Sampson et al., 2015; Wing et al., 2017; Wing et al., 2018). LISFLOOD-FP is a subgrid channel and two-dimensional floodplain model for simulating inundation over large and data sparse areas. The flood hazard map in the CONUS was first simulated by Wing et al. (2017) using 1″ (~30 m) resolution US National Elevation Dataset (NED) terrain information. The design floods in the original simulations using this framework were generated by the index flood method of Smith et al. (2015) and did not consider dam effects.

6.1. Dam Module

The flood routing process considering dam effects has been developed as a new module in LISFLOOD-FP and is shown in Figure 8. The upstream boundaries (QU) are generated using the index flood method and the rational equation. Without a dam, the design flood routes from the upstream boundaries (time varying flow) to the downstream boundary (fixed free surface elevation). In theory, the discharge should increase from upstream to downstream as flood water from tributaries flows into the main reach. When considering the impact of reservoir operations, there should be a reduction in the discharge downstream of dam (Figure 8 (a)). In the enhanced hydrodynamic model, a new internal boundary (QO) is generated at the dam location based on the dam inflow (QI) and the relevant FAI value (Figure 8 (b)). The flood hazard of the dam downstream is then simulated with new internal boundary (QO) and a downstream free surface elevation boundary (HD).

An ideal case study was created to test the effectiveness of the dam module in the LISFLOOD-FP model. The case study is shown in Figure 8 (c) and consists of a simple terrain decrease from the upstream to downstream. We found that the FAI is a sensitive parameter for downstream flood hazard. When FAI is set to 1, the dam inflow is equal to the dam outflow, and the flood hazard downstream of the dam will not change. The flood hazard downstream of the dam is reduced as FAI decreases from 1 to 0.5 to 0.1, which reveals the effectiveness of the dam module in flood hazard mapping.

6.2. Map Comparison

To further validate the reliability of the FAI value in flood hazard mapping, two real case studies in the CONUS were selected: Roy Ink dam (C1) and Mansfield dam (C2) in the Colorado river basin. The

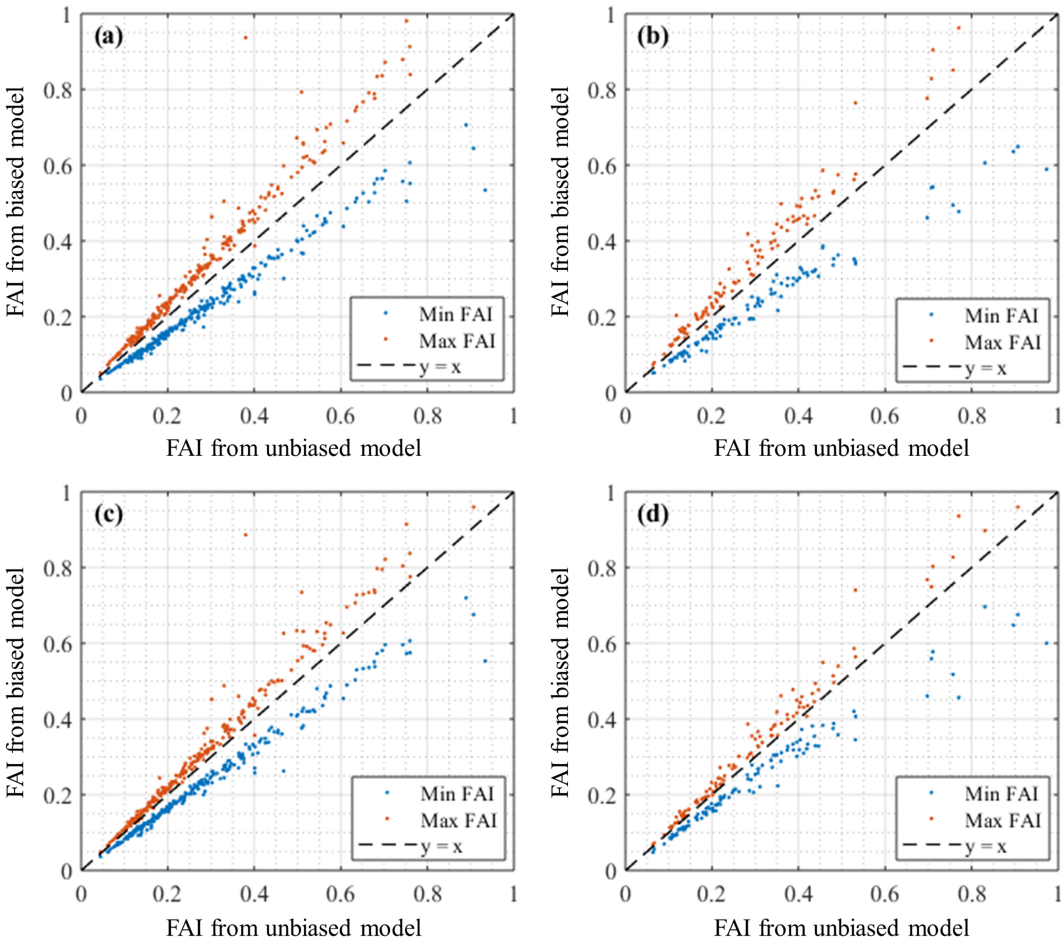

Figure 7. The FAI range of the selected dams: A QI with ±20% bias in training; b QI with±20% bias in testing; c QO with ±20% bias in training; d QO with ±20% bias in testing.

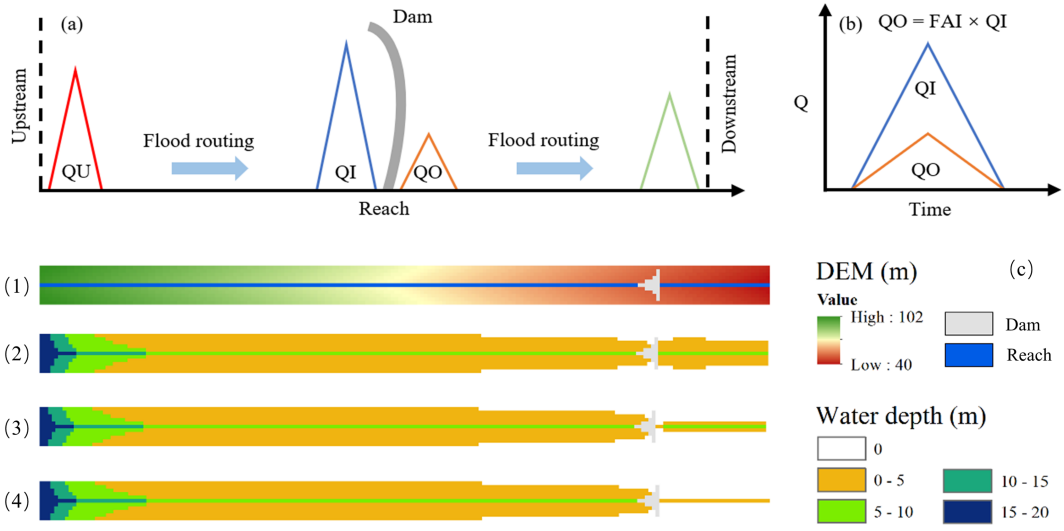

Figure 8. The new LISFLOOD-FP dam module: (a) flood routing process considering dam impact; (b) QO generation process. (c) Testing of dam module in LISFLOOD-FP model: (1) case study; (2) FAI = 1;(3) FAI = 0.5; (4) FAI = 0.1.

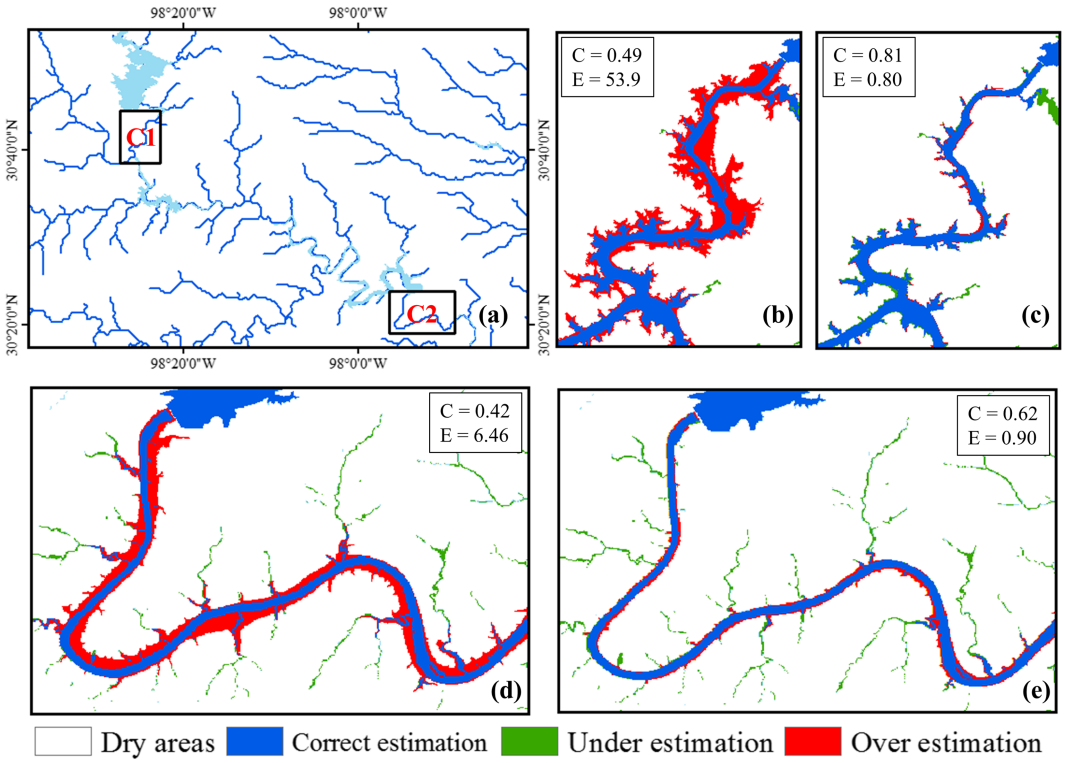

Figure 9. Flood hazard map of the two case studies: A locations of C1 and C2; b map comparison for C1 where FAI = 1; c map comparison for C1 where FAI = 0.22; d map comparison for C2 where FAI = 1; e map comparison for C2 where FAI = 0.13.

locations of the case studies are shown in Figure 9 (a). The Flood Attenuation Index values from the RF model for Roy Ink dam and Mansfield dam were 0.22 and 0.13 respectively. The simulated flood hazard using these FAI values and the new dam module for C1 and C2 were compared with FEMA maps with and without the dam effect (Figure 9 (b-e)). The Critical Success Index for Roy Ink dam increased from 0.49 to 0.81 and for Mansfield dam from 0.42 to 0.62. This showed that the model performance in the two case studies is significantly enhanced by considering dam effects in flood hazard mapping. There is an obvious overestimation of flood extent downstream of each dam when reservoir operations are ignored. These overestimations were also evaluated by the Error Bias metric. Error Bias reduces from 53.9 to 0.80 for Roy Ink dam and from 6.46 to 0.9 for Mansfield dam when using the appropriate FAI value in flood hazard mapping. This shows that overestimation in flood hazard maps caused by not taking dams into account can be significantly corrected using the proposed method.

7. Conclusions

A Random Forest based approach was proposed to estimate the impact of dams on design flood magnitudes based on public datasets. As a proof of concept, the analysis is performed for the 100-year return period flood but could easily be extended to a range of return periods. This approach can provide an effective strategy for inferring the flood attenuation effect of reservoirs in ungauged areas and improve flood hazard map accuracy when coupled with hydraulic models. At present the method has been demonstrated over the CONUS but could be extended globally. In fact, the wide range of hydroclimates present across the US means that the results may already be quite widely applicable.

The main conclusions of this study are summarized below:

1. The RF-based method successfully characterized the relationship between the impact of dams on the design flood and explanatory factors contained in publicly available datasets. The RF-based approach showed a higher accuracy in terms of the Nash Sutcliffe efficiency coefficient (0.92 in training and 0.87

in testing) and outperformed an alternative Multi-Linear Regression-based approach (0.68 in training and 0.61 in testing).

2. The contribution of explanatory factors to dam outflow was evaluated with both RF and CFS models. It was shown that streamflow characteristics of the basin (long-term average discharge, area of upstream catchment and dam inflow) were the most important predictors of dam outflow for the 100-year return period flow. For areas that dam inflow is not available, the RF model can provide a satisfactory prediction of the attenuation of the 100-year return period flow based on the parameters in GRanD database.
3. The flood attenuation index (FAI) of dams for 100-year return period flood flows over the CONUS were simulated with the proposed method. There was no obvious pattern or trend in FAI at the spatial scale of the CONUS, and each large river had both high and low attenuation dams. About half of dams (47%) have high flood attenuation effects (FAI <0.4) and 36% dams have a low flood attenuation effects (FAI >0.6).
4. The flood routing process considering FAI has been developed as a new module in the LISFLOOD-FP two-dimensional hydrodynamic model. Two case studies in the Colorado river basin were selected to test the accuracy of FAI and the enhanced LISFLOOD-FP model. There is an obvious overestimation of flood extent downstream of each dam when reservoir operations are not taken into account. These overestimations can be significantly corrected using the proposed approach. Whilst developed for the LISFLOOD-FP code the proposed method is generic and could be easily implemented in any other hydrodynamic model.
5. The impact of dam in/outflow uncertainty on RF model accuracy and FAI values was quantitatively evaluated. Dam inflow uncertainty has less impact on RF accuracy and the RF model can provide a reliable prediction (NSE > 0.85 during testing) when randomly adding ±20% bias. The proposed RF model will not exacerbate the effect of discharge uncertainty on FAI, and dam inflow (QI) uncertainty has larger impact on the FAI value than uncertainties in dam outflow (QO). Although some individual large error points were found under both dam in/outflow uncertainty, the proposed method could nevertheless provide valuable information for decision making at the national scale and is certainly a significant improvement over current large scale methods (e.g. Wing et al., 2017) which do not account for flood attenuation by dams at all.

Acknowledgments

Gang Zhao would like to thank Oliver Wing, Andrew Smith and Christopher Sampson at Fathom for helpful discussions. Paul Bates is supported by a Royal Society Research Merit award. Jeff Neal is supported by Natural Environment Research Council (NERC) grants (NE/S003061/1 and NE/S006079/1). Gang Zhao is supported by the China Scholarship Council-University of Bristol Joint PhD Scholarships Programme (No.201700260088). FEMA flood maps are available from the FEMA Flood Map Service Center (https://msc.fema.gov/portal/home). Global Reservoir and Dam (GRanD) database is available from (http://sedac.ciesin.columbia.edu/data/set/grand-v1-dams-rev01). To access discharge data used in extreme flood estimation, visit https://water-data.usgs.gov/nwis for USGS surface-water data. The terrain data used in case studies is available from US National Elevation Dataset (NED) (https://catalog.data.gov/dataset/usgs--national-elevation-dataset-ned). The LISFLOOD-FP model is available for non-commercial research purposes by contacting Paul Bates at the University of Bristol (http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/). The Random Forest model is implemented based on '*randomForest*' package (https://cran.r-project.org/web/packages/randomForest/randomForest.pdf).

References

Ahmad, A., El-Shafie, A., Razali, S. F. M., & Mohamad, Z. S. (2014). Reservoir optimization in water resources: A review. *Water Resources Management*, *28*(11), 3391–3405. https://doi.org/10.1007/s11269-014-0700-5

Ashley, R. M., Balmforth, D. J., Saul, A. J., & Blanskby, J. D. (2005). Flooding in the future - predicting climate change, risks and responses in urban areas. *Water Science and Technology*, *52*(5), 265–273.

Assani, A. A., Stichelbout, E., Roy, A. G., & Petit, F. (2006). Comparison of impacts of dams on the annual maximum flow characteristics in three regulated hydrologic regimes in Québec (Canada). *Hydrological Processes: An International Journal*, *20*(16), 3485–3501.

Biemans, H., Haddeland, I., Kabat, P., Ludwig, F., Hutjes, R., Heinke, J., & Gerten, D. (2011). Impact of reservoirs on river discharge and irrigation water supply during the 20th century. *Water Resources Research*, *47*, W03509. https://doi.org/10.1029/2009WR008929

Bocchiola, D., De Michele, C., & Rosso, R. (2003). Review of recent advances in index flood estimation. *Hydrology and Earth System Sciences Discussions*, *7*(3), 283–296.

Breiman, L. (2002). Manual on setting up, using, and understanding random forests v3. 1. *Statistics Department University of California Berkeley, CA, USA, 1*, 58.

Changnon, S. A. (2008). Assessment of flood losses in the United States. *Journal of Contemporary Water Research & Education*, *138*(1), 38–44.

Dalrymple, T. (1960). Flood-frequency analyses, manual of hydrology: Part 3. Retrieved from

Döll, P., Fiedler, K., & Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences*, *13*(12), 2413.

Erskine, W. D., Terrazzolo, N., & Warner, R. (1999). River rehabilitation from the hydrogeomorphic impacts of a large hydro-electric power project: Snowy River, Australia. *Regulated Rivers: Research & Management: An International Journal Devoted to River Research and Management*, *15*(1–3), 3–24.

FitzHugh, T. W., & Vogel, R. M. (2011). The impact of dams on flood flows in the United States. *River Research and Applications*, *27*(10), 1192–1215.

Fortier, C., Assani, A. A., Mesfioui, M., & Roy, A. G. (2011). Comparison of the interannual and interdecadal variability of heavy flood characteristics upstream and downstream from dams in inversed hydrologic regime: Case study of Matawin River (Québec, Canada). *River Research and Applications*, *27*(10), 1277–1289.

Friesen, J., Andreini, M., Andah, W., Amisigo, M. A., & Giesen, N. V. D. (2005). Storage capacity and long-term water balance of the Volta basin, West Africa. In *Proceedings of symposium S6 held during the Seventh IAHS Scientific Assembly*. Brazil and UK: IAHS Publication. Retrieved from http://iahs.info/uploads/dms/13203.2220138-14520Foz20S6-2-1420Friesen.pdf

Gleick, P. H. (2012). China dams. In *The world's water* (pp. 127–142). Washington, DC: Island Press.

Graf, W. L. (1999). Dam nation: A geographic census of American dams and their large-scale hydrologic impacts. *Water Resources Research*, *35*(4), 1305–1311.

Guo, S., Chen, J., Li, Y., Liu, P., & Li, T. (2011). Joint operation of the multi-reservoir system of the three gorges and the Qingjiang cascade reservoirs. *Energies*, *4*(7), 1036–1050.
Hall, M. A. (1999). Correlation-Based Feature Selection for Machine Learning.
Hanasaki, N., Kanae, S., & Oki, T. (2006). A reservoir operation scheme for global river routing models. *Journal of Hydrology*, *327*(1–2), 22–41.
Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276
Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., & Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816.
ICOLD. (2018). World register of dams: International Commission on Large Dams.
Jha, A. K., Bloch, R., & Lamond, J. (2012). *Cities and flooding: a guide to integrated urban flood risk management for the 21st century*. Washinton, DC: The World Bank.
Kjeldsen, T. R., & Jones, D. A. (2006). Prediction uncertainty in a median-based index flood method using L moments. *Water Resources Research*, *42*, W07414. https://doi.org/10.1029/2005WR004069
Kjeldsen, T. R., Smithers, J., & Schulze, R. (2002). Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. *Journal of Hydrology*, *255*(1–4), 194–211.
Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Magome, J. (2011). Global reservoir and dam (grand) database. Technical Documentation, Version, 1.
Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94.
Linnerooth-Bayer, J., & Mechler, R. (2015). Insurance for assisting adaptation to climate change in developing countries: A proposed strategy. *Climate Policy*, *6*(6), 621-636. https://doi.org/10.1080/14693062.2006.968562
Lloyd, S. P. (1982). Least-squares quantization in Pcm. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/Tit.1982.1056489
McMahon, T. A., Pegram, G. G. S., Vogel, R. M., & Peel, M. C. (2007). Revisiting reservoir storage-yield relationships using a global streamflow database. *Advances in Water Resources*, *30*(8), 1858–1872. https://doi.org/10.1016/j.advwatres.2007.02.003
Mei, X., Van Gelder, P., Dai, Z., & Tang, Z. (2017). Impact of dams on flood occurrence of selected rivers in the United States. *Frontiers of Earth Science*, *11*(2), 268–282.
Ntelekos, A. A., Oppenheimer, M., Smith, J. A., & Miller, A. J. (2010). Urbanization, climate change and flood policy in the United States. *Climatic Change*, *103*(3–4), 597–616.
Pan, L., Housh, M., Liu, P., Cai, X., & Chen, X. (2015). Robust stochastic optimization for reservoir operation. *Water Resources Research*, *51*, 409–429. https://doi.org/10.1002/2014WR015380
Pang, B., Yue, J., Zhao, G., & Xu, Z. (2017). Statistical downscaling of temperature with the random Forest model. *Advances in Meteorology*, *2017*, 1–11. https://doi.org/10.1155/2017/7265178
Pappenberger, F., Dutra, E., Wetterhall, F., & Cloke, H. L. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade. *Hydrology and Earth System Sciences*, *16*(11), 4143–4156.
Richter, B. D., Baumgartner, J. V., Braun, D. P., & Powell, J. (1998). A spatial assessment of hydrologic alteration within a river network. *Regulated Rivers: Research & Management: An International Journal Devoted to River Research and Management*, *14*(4), 329–340.
Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., & Freer, J. E. (2015). A high-resolution global flood hazard model. *Water Resources Research*, *51*(9), 7358–7381. https://doi.org/10.1002/2015WR016954
Seaber, P. R., Kapinos, F. P., & Knapp, G. L. (1987). *Hydrologic map units*. Denver: United States Government Printing Office.
Smith, A., Sampson, C., & Bates, P. (2015). Regional flood frequency analysis at the global scale. *Water Resources Research*, *51*, 539–553. https://doi.org/10.1002/2014WR015814
Valdes, J. B., & Marco, J. B. (1995). Managing reservoirs for flood control. Paper Presented at the Proceedings of the US–Italy Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods, Perugia, Italy.
Wang, Z. L., Lai, C. G., Chen, X. H., Yang, B., Zhao, S. W., & Bai, X. Y. (2015). Flood hazard risk assessment model based on random forest. *Journal of Hydrology*, *527*, 1130–1141. Retrieved from <go to ISI>://WOS:000358629100095
Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236. https://doi.org/10.2307/2282967
Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., van Beek, R., Bierkens, M. F., et al. (2013). Assessing flood risk at the global scale: Model setup, results, and sensitivity. *Environmental Research Letters*, *8*(4), 044019. https://doi.org/10.1088/1748-9326/8/4/044019
Wing, O. E., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erickson, T. A. (2017). Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resources Research*, *53*, 7968–7986. https://doi.org/10.1002/2017WR020917
Wing, O. E., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Fargione, J., & Morefield, P. (2018). Estimates of present and future flood risk in the conterminous United States. *Environmental Research Letters*, *13*(3), 034023.
Woznicki, S. A., Baynes, J., Panlasigui, S., Mehaffey, M., & Neale, A. (2019). Development of a spatially complete floodplain map of the conterminous United States using random forest. *Science of the Total Environment*, *647*, 942–953.
Yang, G., Guo, S., Liu, P., Li, L., & Xu, C. (2017). Multiobjective reservoir operating rules based on cascade reservoir input variable selection method. *Water Resources Research*, *53*, 3446–3463. https://doi.org/10.1002/2016WR020301
Zajac, Z., Revilla-Romero, B., Salamon, P., Burek, P., Hirpa, F. A., & Beck, H. (2017). The impact of lake and reservoir parameterization on global streamflow simulation. *Journal of Hydrology*, *548*, 552–568. https://doi.org/10.1016/j.jhydrol.2017.03.022
Zhao, G., Pang, B., Xu, Z. X., Yue, J. J., & Tu, T. B. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. *Science of the Total Environment*, *615*, 1133–1142.
Zhou, T., Nijssen, B., Gao, H., & Lettenmaier, D. P. (2016). The contribution of reservoirs to global land surface water storage variations. *Journal of Hydrometeorology*, *17*(1), 309–325.